Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity

International Symposium on the Art of Statistical Metaware
March 14—16, 2005, Tokyo, Japan

Hayato Waki
Tokyo Institute of Technology, Tokyo, Japan

Sunyoung Kim
Ewha Women's University, Seoul, Korea

Masakazu Kojima
Tokyo Institute of Technology, Tokyo, Japan

Masakazu Muramatsu
The University of Electro-Communications, Tokyo, Japan

Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity

International Symposium on the Art of Statistical Metaware
March 14—16, 2005, Tokyo, Japan

- **An introduction to the recent development of SOS and SDP relaxations for computing global optimal solutions of POPs**

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

- Sparsity and Numerical results are main contributions of the paper.

Outline

1. **POPs (Polynomial Optimization Problems)**
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

$\mathbb{R}^n$: the n-dim Euclidean space.

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$: a vector variable.

$f_j(x)$: a multivariate polynomial in $x \in \mathbb{R}^n$ $(j = 0, 1, \ldots, m)$.

POP (Poly. Opt. Prob.): min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

$\mathbb{R}^n$: the n-dim Euclidean space.

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$: a vector variable.

$f_j(x)$: a multivariate polynomial in $x \in \mathbb{R}^n$ $(j = 0, 1, \ldots, m)$.

POP (Poly. Opt. Prob.): min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

Example: $n = 3$

$$
\begin{array}{rl}
\min & f_0(x) \equiv x_1^3 - 2x_1x_2^2 + x_1^2x_2x_3 - 4x_3^2 \\
\text{sub.to} & f_1(x) \equiv -x_1^2 + 5x_2x_3 + 1 \geq 0, \\
 & f_2(x) \equiv x_1^2 - 3x_1x_2x_3 + 2x_3 + 2 \geq 0, \\
 & f_3(x) \equiv -x_1^2 - x_2^2 - x_3^2 + 1 \geq 0.
\end{array}
$$

$\mathbb{R}^n$: the n-dim Euclidean space.

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$: a vector variable.

$f_j(x)$: a multivariate polynomials in $x \in \mathbb{R}^n$ $(j = 0, 1, \ldots, m)$.

POP (Poly. Opt. Prob.): min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

Example: $n = 3$

$$\begin{array}{rl} \min & f_0(x) \equiv x_1^3 - 2x_1x_2^2 + x_1^2x_2x_3 - 4x_3^2 \\ \text{sub.to} & f_1(x) \equiv -x_1^2 + 5x_2x_3 + 1 \geq 0, \\ & f_2(x) \equiv x_1^2 - 3x_1x_2x_3 + 2x_3 + 2 \geq 0, \\ & f_3(x) \equiv -x_1^2 - x_2^2 - x_3^2 + 1 \geq 0, \\ & x_1(x_1 - 1) = 0 \text{ (0-1 integer)}, \\ & x_2 \geq 0, \; x_3 \geq 0, \; x_2x_3 = 0 \text{ (complementarity)}. \end{array}$$

- Various problems can be described as POPs.
- A unified theoretical model for global optimization in nonlinear and combinatorial optimization problems.

Two approaches to SOS and SDP relaxations of POPs

POP: $\min\ f_0(x)$ sub.to $f_i(x) \geq 0\ (i = 1, \ldots, m),$

POP	$\Rightarrow$	generalized Lagrangian dual
$\Updownarrow$ add valid LMIs	dual	$\Downarrow$
Polynomial SDP		$\Downarrow$ SOS relaxation
$\Downarrow$ linearize (relaxation)	dual	$\Downarrow$
SDP[1]	$\Leftrightarrow$	SDP[2]

[1] J.B.Lasserre, "Global optimization with polynomials and the problems of moments", *SIAM J. on Optimization*, 11 (2001) 796–817.

[2] P.A.Parrilo, "Semidefinite programming relaxations for semialgebraic problems". *Math. Prog.*, 96 (2003) 293–320.

Two approaches to SOS and SDP relaxations of POPs

$$\textbf{POP:} \quad \min \ f_0(x) \quad \text{sub.to} \quad f_i(x) \geq 0 \ (i = 1, \ldots, m),$$

POP	$\Rightarrow$	generalized Lagrangian dual
$\Updownarrow$ add valid LMIs	dual	$\Downarrow$
Polynomial SDP		$\Downarrow$ SOS relaxation
$\Downarrow$ linearize (relaxation)	dual	$\Downarrow$
SDP[1]	$\Leftrightarrow$	SDP[2]

[1] J.B.Lasserre, "Global optimization with polynomials and the problems of moments", *SIAM J. on Optimization*, 11 (2001) 796–817.

[2] P.A.Parrilo, "Semidefinite programming relaxations for semialgebraic problems". *Math. Prog.*, 96 (2003) 293–320.

(a) Global optimal solutions.

(b) Large-scale SDPs require enormous computation.

(c) Proposed SDP relaxation = SDP[1] + "Exploiting sparsity".

Outline

1. POPs (Polynomial Optimization Problems)
2. **A sequence of relaxations**
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

$$\boldsymbol{\mathcal{P}} \textbf{ (POP): min } \boldsymbol{f_0(x)} \textbf{ sub.to } \boldsymbol{x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}.}$$

$\mathcal{P}$ (POP): min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$.

$\Longrightarrow$ **A sequence $\{\mathcal{P}^r\}$ of relaxations of $\mathcal{P}$ with increasing size:**

(a) Each $\mathcal{P}^r$ is a convex program (SDP), and can be solved numerically.

(b) opt.val. of $\mathcal{P}^r \leq$ opt.val. of $\mathcal{P}^{r+1} \leq$ opt.val. of $\mathcal{P}$.

(c) In practice, opt.val. of $\mathcal{P}^r =$ opt.val. of $\mathcal{P}$ for some small r.

$\mathcal{P}$ (POP): min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$.

$\Longrightarrow$ A sequence $\{\mathcal{P}^r\}$ of relaxations of $\mathcal{P}$ with increasing size:

(a) Each $\mathcal{P}^r$ is a convex program (SDP), and can be solved numerically.

(b) opt.val. of $\mathcal{P}^r \leq$ opt.val. of $\mathcal{P}^{r+1} \leq$ opt.val. of $\mathcal{P}$.

(c) In practice, opt.val. of $\mathcal{P}^r =$ opt.val. of $\mathcal{P}$ for some small r.

ex3_1_4 from globallib: 3 variables and 9 consraints, opt.val. $= -4.000$.

$\{\mathcal{P}^r\}$	m	size(A_i)	# nonzeros in A_i's	lower bound	cpu
$\mathcal{P}^1$	9	25×25	47	-6.000	0.21
$\mathcal{P}^2$	34	108×108	571	-5.591	0.75
$\mathcal{P}^3$	84	270×270	3153	-4.062	0.81
$\mathcal{P}^4$	164	537×537	11940	-4.000	2.04

- Each SDP $\mathcal{P}^r$ has the form: min $\sum_{i=1}^{m} b_i y_i$ sub.to $\sum_{i=1}^{m} A_i y_i - A_0 \succeq O$.

$\mathcal{P}$ (POP): min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$.

$\Longrightarrow$ A sequence $\{\mathcal{P}^r\}$ of relaxations of $\mathcal{P}$ with increasing size:

(a) Each $\mathcal{P}^r$ is a convex program (SDP), and can be solved numerically.

(b) opt.val. of $\mathcal{P}^r \leq$ opt.val. of $\mathcal{P}^{r+1} \leq$ opt.val. of $\mathcal{P}$.

(c) In practice, opt.val. of $\mathcal{P}^k$ = opt.val. of $\mathcal{P}$ for some small k.

ex3_1_4 from globallib: 3 variables and 9 consraints, opt.val. $= -4.000$.

$\{\mathcal{P}^r\}$	m	size(A_i)	# nonzeros in A_i's	lower bound	cpu
$\mathcal{P}^1$	9	25×25	47	-6.000	0.21
$\mathcal{P}^2$	34	108×108	571	-5.591	0.75
$\mathcal{P}^3$	84	270×270	3153	-4.062	0.81
$\mathcal{P}^4$	164	537×537	11940	-4.000	2.04

• Each SDP $\mathcal{P}^r$ has the form: min $\sum_{i=1}^{m} b_i y_i$ sub.to $\sum_{i=1}^{m} A_i y_i - A_0 \succeq O$.

• The size of $\mathcal{P}^r$ gets larger rapidly.

• To solve larger POPs,
"how to exploit the sparsity in polynomials and SDPs" is a key issue.

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

$f(x)$: a nonnegative polynomial in $x \in \mathbb{R}^n$ $\Leftrightarrow$ $f(x) \geq 0$ $(\forall x \in \mathbb{R}^n)$.

$\mathcal{N}$: the set of nonnegative polynomials in $x \in \mathbb{R}^n$.

$f(x)$: a nonnegative polynomial in $x \in \mathbb{R}^n \Leftrightarrow f(x) \geq 0 \ (\forall x \in \mathbb{R}^n)$.

$\mathcal{N}$: the set of nonnegative polynomials in $x \in \mathbb{R}^n$.

$f(x)$: an SOS (Sum of Squares) polynomial

$$\Updownarrow$$

$\exists$ a finite number of polynomials $g_1(x), \ldots, g_k(x)$; $f(x) = \sum_{i=1}^{k} g_i(x)^2$.

SOS_* : the set of SOS. Obviously, $\mathrm{SOS}_* \subset \mathcal{N}$.

$\mathrm{SOS}_{2r} = \{f \in \mathrm{SOS}_* : \deg f \leq 2r\}$: the set of SOS with degree ar most $2r$.

$n = 2$. $f(x_1, x_2) = (x_1^2 - 2x_2 + 1)^2 + (3x_1x_2 + x_2 - 4)^2 \in \mathrm{SOS}_4$.

$n = 2$. $f(x_1, x_2) = (x_1x_2 - 1)^2 + x_1^2 \in \mathrm{SOS}_4$.

$f(x)$: a nonnegative polynomial in $x \in \mathbb{R}^n \quad \Leftrightarrow \quad f(x) \geq 0 \ (\forall x \in \mathbb{R}^n)$.

$\mathcal{N}$: the set of nonnegative polynomials in $x \in \mathbb{R}^n$.

$f(x)$: an SOS (Sum of Squares) polynomial

$\Updownarrow$

$\exists$ a finite number of polynomials $g_1(x), \ldots, g_k(x)$; $f(x) = \sum_{i=1}^{k} g_i(x)^2$.

SOS_* : the set of SOS. Obviously, $\mathrm{SOS}_* \subset \mathcal{N}$.

$\mathrm{SOS}_{2r} = \{f \in \mathrm{SOS}_* : \deg f \leq 2r\}$: the set of SOS with degree ar most $2r$.

- In theory, SOS_* (SOS) $\subset \mathcal{N}$ (nonnegative).
- If $n = 1$, $\mathrm{SOS}_* = \mathcal{N}$. $\{f \in \mathcal{N} : \deg f \leq 2\} \equiv \mathrm{SOS}_2$. $\mathrm{SOS}_* \neq \mathcal{N}$ in general.
- In practice, $f(x) \in \mathcal{N} \setminus \mathrm{SOS}_*$ is rare.
- So we replace $\mathcal{N}$ by $\mathrm{SOS}_* \Longrightarrow$ SOS Relaxations.

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. **SOS relaxation of unconstrained POPs**
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

$$\mathcal{P}\text{:} \quad \min_{x \in \mathbb{R}^n} \; f(x), \text{ where } f \text{ is a polynomial in } x \in \mathbb{R}^n \text{ with } \deg f = 2r$$

$$\mathcal{P}: \quad \min_{x \in \mathbb{R}^n} f(x), \text{ where } f \text{ is a polynomial in } x \in \mathbb{R}^n \text{ with } \deg f = 2r$$

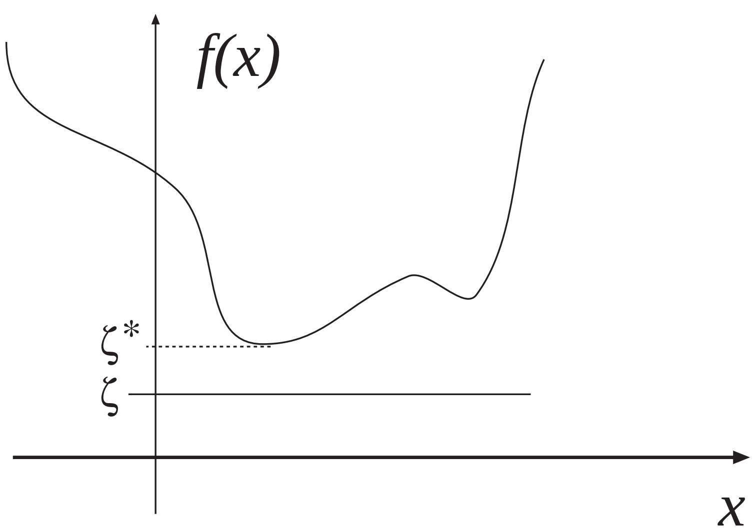
$$\Updownarrow$$

$$\mathcal{P}': \ \max \ \zeta \ \text{ s.t } \quad f(x) - \zeta \geq 0 \ (\forall x \in \mathbb{R}^n)$$

$$\Updownarrow$$

$$f(x) - \zeta \in \mathcal{N} \text{ (the nonnegative polynomials in } x \in \mathbb{R}^n \text{)}$$

Here x is an index describing an infinite number of inequality constraints.



$$\mathcal{P}\text{:} \quad \min_{x \in \mathbb{R}^n} \; f(x), \text{ where } f \text{ is a polynomial in } x \in \mathbb{R}^n \text{ with deg } f = 2r$$

$$\Updownarrow$$

$$\mathcal{P}\text{': max } \zeta \text{ s.t } \quad f(x) - \zeta \geq 0 \ (\forall x \in \mathbb{R}^n)$$

$$\Updownarrow$$

$$f(x) - \zeta \in \mathcal{N} \text{ (the nonnegative polynomials in } x \in \mathbb{R}^n \text{)}$$

Here x is an index describing an infinite number of inequality constraints.

$$\Sigma \subset \text{SOS}_{2r} \subset \text{SOS}_* \subset \mathcal{N} \ \Downarrow \text{ a subproblem of } \mathcal{P}' = \text{a relaxation of } \mathcal{P}$$

$$\mathcal{P}\text{'': max } \zeta \text{ sub.to } f(x) - \zeta \in \Sigma$$

$\mathcal{P}$: $\min_{x \in \mathbb{R}^n} f(x)$, where f is a polynomial in $x \in \mathbb{R}^n$ with deg $f = 2r$

$$\Updownarrow$$

$\mathcal{P}$': max ζ s.t $f(x) - \zeta \geq 0$ $(\forall x \in \mathbb{R}^n)$

$$\Updownarrow$$

$f(x) - \zeta \in \mathcal{N}$ (the nonnegative polynomials in $x \in \mathbb{R}^n$)

Here x is an index describing an infinite number of inequality constraints.

$\Sigma \subset \mathrm{SOS}_{2r} \subset \mathrm{SOS}_* \subset \mathcal{N}$ $\Downarrow$ a subproblem of $\mathcal{P}'$ = a relaxation of $\mathcal{P}$

$\mathcal{P}$": max ζ sub.to $f(x) - \zeta \in \Sigma$

- the min. value of $\mathcal{P}$ = the max. value of $\mathcal{P}'$ $\geq$ the max. value of $\mathcal{P}$"
- $\mathcal{P}$" can be solved as an SDP.
- We can exploit the sparsity of the Hessian matrix of f to reduce the size of Σ; hence the size of the resulting SDP.

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. **SOS relaxation of constrained POPs**
6. Sparsity
7. Numerical results
8. Concluding remarks

- **This part is a little bit complicated!**

“(Generalized) Lagrangian Dual”

$+$

“SOS relaxation of unconstrained POPs”

$\Downarrow$

SOS relaxation of constrained POPs

POP: min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

Generalized Lagrange function:

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) \ \cdots \ - \varphi_m(x) f_m(x).$$

where, $\varphi \in \mathrm{SOS}_*^m \equiv \{\varphi = (\varphi_1, \ldots, \varphi_m) : \varphi_j \in \mathrm{SOS}_*$ **(SOS polynomials)**$\}$.

POP: $\min \ f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

Generalized Lagrange function:

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) \ \cdots \ - \varphi_m(x) f_m(x).$$

where, $\varphi \in \mathrm{SOS}_*^m \equiv \{\varphi = (\varphi_1, \ldots, \varphi_m) : \varphi_j \in \mathrm{SOS}_*$ (SOS polynomials)$\}$.

G. Lagrange relaxation: Given a $\varphi \in \mathrm{SOS}_*^m$, $\min_{x \in \mathbb{R}^n} L(x, \varphi)$.

$$\min_{x \in \mathbb{R}^n} L(x, \varphi) \leq \min_{x \in S} f_0(x) \text{ for } \forall \varphi \in \mathrm{SOS}_*^m.$$

POP: $\min \ f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

Generalized Lagrange function:

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) \ \cdots \ - \varphi_m(x) f_m(x).$$

where, $\varphi \in \mathrm{SOS}_*^m \equiv \{\varphi = (\varphi_1, \ldots, \varphi_m) : \varphi_j \in \mathrm{SOS}_* \text{ (SOS polynomials)}\}$.

G. Lagrange relaxation: Given a $\varphi \in \mathrm{SOS}_*^m$, $\min_{x \in \mathbb{R}^n} L(x, \varphi)$.

$$\min_{x \in \mathbb{R}^n} L(x, \varphi) \leq \min_{x \in S} f_0(x) \text{ for } \forall \varphi \in \mathrm{SOS}_*^m.$$

G. Lagrange dual (the best G.L. relaxation): $\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi)$.

$$\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi) \leq \min_{x \in S} f_0(x).$$

- Under appropriate assumptions, $\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi) = \min_{x \in S} f_0(x)$.

POP: $\min\ f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0\ (j = 1, \ldots, m)\}$

$L(x, \varphi)$: generalized Lagrange function

G. Lagrange dual: $\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi)$

POP: min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

$L(x, \varphi)$: generalized Lagrange function

G. Lagrange dual: $\max_{\varphi \in \mathbf{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi)$

$$\Updownarrow$$

G. Lagrange dual: max ζ s.t $L(x, \varphi) - \zeta \geq 0 \ (\forall x \in \mathbb{R}^n), \ \varphi \in \mathbf{SOS}_*^m$

POP: min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

$L(x, \varphi)$: generalized Lagrange function

G. Lagrange dual: $\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi)$

$\Updownarrow$

G. Lagrange dual: max ζ s.t $L(x, \varphi) - \zeta \geq 0 \ (\forall x \in \mathbb{R}^n)$, $\varphi \in \mathrm{SOS}_*^m$

SOS relaxation $\Downarrow$

max ζ s.t $L(x, \varphi) - \zeta \in \mathrm{SOS}_*$, $\varphi \in \mathrm{SOS}_*^m$

$$\textbf{POP:}\ \min\ f_0(x)\ \text{sub.to}\ x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0\ (j = 1, \ldots, m)\}$$

$L(x, \varphi)$: generalized Lagrange function

$$\textbf{G. Lagrange dual:}\ \max_{\varphi \in \mathrm{SOS}_*^m}\ \min_{x \in \mathbb{R}^n}\ L(x, \varphi)$$

$$\Updownarrow$$

$$\textbf{G. Lagrange dual:}\ \max\ \zeta\ \text{s.t}\ L(x, \varphi) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n),\ \varphi \in \mathrm{SOS}_*^m$$

SOS relaxation $\Downarrow$

$$\max\ \zeta\ \text{s.t}\ L(x, \varphi) - \zeta \in \mathrm{SOS}_*,\ \varphi \in \mathrm{SOS}_*^m$$

a finite size $\Downarrow$ $\Xi \subset \{\varphi(x) = (\varphi_1, \ldots, \varphi_m) : \varphi_j \in \mathrm{SOS}_{2r}\}$ for $\exists r$,
$\Sigma \subset \mathrm{SOS}_{2s}$ for $\exists s \geq r$

$$\textbf{SOS relaxation:}\ \max\ \zeta\ \text{s.t}\ L(x, \varphi) - \zeta \in \Sigma,\ \varphi \in \Xi$$

POP: $\min\ f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0\ (j = 1, \ldots, m)\}$

$L(x, \varphi)$: generalized Lagrange function

G. Lagrange dual: $\displaystyle\max_{\varphi \in \mathrm{SOS}_*^m} \min_{x \in \mathbb{R}^n} L(x, \varphi)$

$\Updownarrow$

G. Lagrange dual: $\max\ \zeta$ s.t $L(x, \varphi) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n)$, $\varphi \in \mathrm{SOS}_*^m$

SOS relaxation $\Downarrow$

$\max\ \zeta$ s.t $L(x, \varphi) - \zeta \in \mathrm{SOS}_*$, $\varphi \in \mathrm{SOS}_*^m$

a finite size $\Downarrow$ $\Xi \subset \{\varphi(x) = (\varphi_1, \ldots, \varphi_m) : \varphi_j \in \mathrm{SOS}_{2r}\}$ for $\exists r$,
$\Sigma \subset \mathrm{SOS}_{2s}$ for $\exists s \geq r$

SOS relaxation: $\max\ \zeta$ s.t $L(x, \varphi) - \zeta \in \Sigma$, $\varphi \in \Xi$

- SOS relaxation can be solved as an SDP.
- As r $\uparrow$, a better lower bound for the opt. val. of POP.
- Sparsity of POP to reduce the sizes of Ξ and Σ.

r : the relaxation order.

Outline

1. POPs (Polynomial Optimization Problems) and their relaxation
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. **Sparsity**
7. Numerical results
8. Concluding remarks

An example of sparse unconstrained POPs — 1 (Conn at el. 1988)

$$f_0(x) = \sum_{j \in J} \left((x_i + 10x_{i+1})^2 + 5(x_{i+2} - x_{i+3})^2 \right.$$
$$\left. +(x_{i+1} - 2x_{i+2})^4 + 10(x_i \ - 10x_{i+3})^4 \right),$$

where $J = \{1, 3, 5, \ldots, n-3\}$ and n is a multiple of 4.

- **The Hessian matrix is sparse (narrow bandwidth).**

An example of sparse unconstrained POPs — 1 (Conn at el. 1988)

$$f_0(x) = \sum_{j \in J} \left((x_i + 10x_{i+1})^2 + 5(x_{i+2} - x_{i+3})^2 \right.$$
$$\left. + (x_{i+1} - 2x_{i+2})^4 + 10(x_i \; - 10x_{i+3})^4 \right),$$

where $J = \{1, 3, 5, \ldots, n-3\}$ and n is a multiple of 4.

- The Hessian matrix is sparse (narrow bandwidth).

Numerical results on sparse and Lasserre's dense relaxations (r=2)

		cpu in sec.	
n	$\epsilon_{\rm obj}$	sparse	Lasserre's dense
12	1.1e-09	0.7	404.2
16	9.0e-10	0.9	7523.1
40	1.7e-09	2.1	—
100	3.6e-04	2.2	—

$$\epsilon_{\rm obj} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}}.$$

An example of sparse unconstrained POPs — 2

Generalized Rosenbrock function (Nash 1984).

$$f_0(x) = 1 + \sum_{i=1}^{n} \left(100(x_i - x_{i-1}^2)^2 + (1 - x_i)^2\right)$$

- **The Hessian matrix is sparse (tridiagonal).**

An example of sparse unconstrained POPs — 2

Generalized Rosenbrock function (Nash 1984).

$$f_0(x) = 1 + \sum_{i=1}^{n} \left(100(x_i - x_{i-1}^2)^2 + (1 - x_i)^2\right)$$

- **The Hessian matrix is sparse (tridiagonal).**

Numerical results on sparse and dense Lasserre's relaxations (r=2)

		cpu in sec.	
n	ϵ_{obj}	sparse	Lasserre's dense
200	1.6e-05	1.8	—
300	3.0e-05	2.5	—
400	1.2e-04	3.3	—
500	4.3e-04	4.5	—

$$\epsilon_{\text{obj}} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}}.$$

An example of sparse constrained POPs: alkyl from globallib

($n = 14$, the max degree of the polynomials in POP $= 3$)

$$
\begin{array}{ll}
\min & -6.3x_4x_7 + 5.04x_1 + 0.35x_2 + x_3 + 3.36x_5 \\
\text{s.t.} & 0.98x_3 - x_6(0.01x_4x_9 + x_3) = 0, \ -x_1x_8 + 10x_2 + x_5 = 0, \\
& x_4x_{11} - x_1(1.12 + 0.13167x_8 - 0.0067x_8x_8) = 0, \\
& \cdots \\
& x_9x_{13} + 22.2x_{10} - 35.82 = 0, \ x_{10}x_{14} - 3x_7 + 1.33 = 0, \\
& \ell_i \leq x_i \leq u_i \ (i = 1, 2, \ldots, 14).
\end{array}
$$

- Each constraints involves only a small number of the variables!

An example of sparse constrained POPs: alkyl from globallib

($n = 14$, the max degree of the polynomials in POP = 3)

$$\begin{array}{ll} \min & -6.3x_4x_7 + 5.04x_1 + 0.35x_2 + x_3 + 3.36x_5 \\ \text{s.t.} & 0.98x_3 - x_6(0.01x_4x_9 + x_3) = 0, \ -x_1x_8 + 10x_2 + x_5 = 0, \\ & x_4x_{11} - x_1(1.12 + 0.13167x_8 - 0.0067x_8x_8) = 0, \\ & \cdots \\ & x_9x_{13} + 22.2x_{10} - 35.82 = 0, \ x_{10}x_{14} - 3x_7 + 1.33 = 0, \\ & \ell_i \leq x_i \leq u_i \ (i = 1, 2, \ldots, 14). \end{array}$$

- Each constraints involves only a small number of the variables!

	sparse			Lasserre's dense		
r (relaxation order)	$\epsilon_{\rm obj}$	$\epsilon_{\rm feas}$	cpu	$\epsilon_{\rm obj}$	$\epsilon_{\rm feas}$	cpu
2	2.0e-03	2.5e-01	6.7	7.3e-06	3.2e-02	65.7
3	9.0e-09	3.0e-08	5216.2	—	—	—

$$\epsilon_{\rm obj} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}},$$

$\epsilon_{\rm feas}$ = the maximum error in the equality constraints.

POP: min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

The basic idea of exploiting sparsity in SOS relaxations:

(a) Choose $\varphi_1(x), \ldots, \varphi_m(x) \in$ SOS such that the sparsity pattern of the Hessian matrix of

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) - \cdots - f_m(x) \varphi_m(x)$$

has a sparse symbolic Cholesky factorization.

POP: min $f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

The basic idea of exploiting sparsity in SOS relaxations:

(a) Choose $\varphi_1(x), \ldots, \varphi_m(x) \in$ SOS such that the sparsity pattern of the Hessian matrix of

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) - \cdots - f_m(x) \varphi_m(x)$$

has a sparse symbolic Cholesky factorization.

(b) For effectiveness of the SOS relaxation, take $\varphi_i(x)$ which involves at least the same set of variables as $f_i(x)$ $(i = 1, 2, \ldots, m)$; for example,

$f_i(x) = 3x_1x_5 + 3x_8^3 \geq 0$
$\Rightarrow \varphi_i(x)$ involves x_1, x_5 and x_8 but not necessarily all other variables.

POP: $\min \ f_0(x)$ sub.to $x \in S \equiv \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, \ldots, m)\}$

The basic idea of exploiting sparsity in SOS relaxations:

(a) Choose $\varphi_1(x), \ldots, \varphi_m(x) \in$ SOS such that the sparsity pattern of the Hessian matrix of

$$L(x, \varphi) = f_0(x) - \varphi_1(x) f_1(x) - \ \cdots \ - f_m(x)\varphi_m(x)$$

has a sparse symbolic Cholesky factorization.

(b) For effectiveness of the SOS relaxation, take $\varphi_i(x)$ which involves at least the same set of variables as $f_i(x)$ $(i = 1, 2, \ldots, m)$; for example,

$f_i(x) = 3x_1x_5 + 3x_8^3 \geq 0$
$\Rightarrow \varphi_i(x)$ involves x_1, x_5 and x_8 but not necessarily all other variables.

POP is correlatively sparse if the sparsity pattern of the Hessian matrix of $L(x, \varphi)$ with any choice of $\varphi_1(x), \ldots, \varphi_m(x) \in$ SOS satisfying (b) has a sparse symbolic Cholesky factorization.

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. **Numerical results**
8. Concluding remarks

Numerical results

Software

- MATLAB for constructing sparse and dense SDP relaxation problems
- SeDuMi to solve SDPs.

Hardware

- 2.4GHz Xeon cpu with 6.0GB memory.

A discrete-time optimal control problem from Coleman et al. 1995

$$\left.\begin{array}{ll} \min & \dfrac{1}{M}\displaystyle\sum_{i=1}^{M-1}\left(y_i^2 + x_i^2\right) \\ \text{s.t.} & y_{i+1} = y_i + \dfrac{1}{M}(y_i^2 - x_i), \quad (i = 1, \ldots, M-1), \quad y_1 = 1. \end{array}\right\}$$

Numerical results on sparse relaxation

M	# of variables	ϵ_{obj}	ϵ_{feas}	cpu
600	1198	3.4e-08	2.2e-10	3.4
700	1398	2.5e-08	8.1e-10	3.3
800	1598	5.9e-08	1.6e-10	3.8
900	1798	1.4e-07	6.8e-10	4.5
1000	1998	6.3e-08	2.7e-10	5.0

$$\epsilon_{\text{obj}} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}},$$

ϵ_{feas} = the maximum error in the equality constraints,

cpu : cpu time in sec. to solve an SDP relaxation problem.

Benchmark problems from globallib

			sparse			Lasserre's dense		
problem	n	r	ϵ_{obj}	ϵ_{feas}	cpu	ϵ_{obj}	ϵ_{feas}	cpu
ex3_1_1	8	3	6.3e-09	6.5e-02	5.5	0.7e-08	2.0e-02	597.8
st_bpaf1b*	10	2	3.8e-08	2.8e-08	1.0	4.6e-09	7.2e-10	1.7
st_e07*	10	2	0.0e+00	8.1e-05	0.4	0.0e+00	8.8e-06	3.0
ex2_1_3	13	2	5.1e-09	3.5e-09	0.5	1.6e-09	1.5e-09	7.7
ex9_1_1	13	2	0.0	4.5e-06	1.5	0.0	9.2e-07	7.7
alkyl*	14	3	9.0e-09	3.0e-08	5216.2	—	—	—
ex9_2_3*	16	2	0.0e+00	5.7e-06	2.3	0.0e+00	7.5e-06	49.7
ex2_1_8*	24	2	1.0e-05	0.0e+00	304.6	3.4e-06	0.0e+00	1946.6

$r =$ relaxation order,

$$\epsilon_{\text{obj}} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}},$$

$\epsilon_{\text{feas}} =$ the maximum error in the equality constraints,

cpu : cpu time in sec. to solve an SDP relaxation problem.

Benchmark problems from globallib

			sparse			Lasserre's dense		
problem	n	r	ϵ_{obj}	ϵ_{feas}	cpu	ϵ_{obj}	ϵ_{feas}	cpu
ex3_1_1	8	3	6.3e-09	6.5e-02	5.5	0.7e-08	2.0e-02	597.8
st_bpaf1b*	10	2	3.8e-08	2.8e-08	1.0	4.6e-09	7.2e-10	1.7
st_e07*	10	2	0.0e+00	8.1e-05	0.4	0.0e+00	8.8e-06	3.0
ex2_1_3	13	2	5.1e-09	3.5e-09	0.5	1.6e-09	1.5e-09	7.7
ex9_1_1	13	2	0.0	4.5e-06	1.5	0.0	9.2e-07	7.7
alkyl*	14	3	9.0e-09	3.0e-08	5216.2	—	—	—
ex9_2_3*	16	2	0.0e+00	5.7e-06	2.3	0.0e+00	7.5e-06	49.7
ex2_1_8*	24	2	1.0e-05	0.0e+00	304.6	3.4e-06	0.0e+00	1946.6

- $\star$ — no tight optimal value before.
- The sparse relaxation attains approx. opt. solutions with the same quality as the dense relaxation.
- The sparse relaxation is much faster than the dense relaxation in large dim. and higher relaxation order cases.

Outline

1. POPs (Polynomial Optimization Problems)
2. A sequence of relaxations
3. Nonnegative polynomials and SOS (Sum of Squares) polynomials
4. SOS relaxation of unconstrained POPs
5. SOS relaxation of constrained POPs
6. Sparsity
7. Numerical results
8. Concluding remarks

- **Lasserre's relaxation**
  — theoretical convergence but expensive in practice.
- **The proposed sparse relaxation**
  = Lasserre's relaxation + sparsity
  — no theoretical convergence but very powerful in practice.

- **Lasserre's dense relaxation**
  — theoretical convergence but expensive in practice.
- **The proposed sparse relaxation**
  = Lasserre's relaxation + sparsity
  — no theoretical convergence but very powerful in practice.
- **There remain many issues to be studied further.**
  - **Exploiting sparsity.**
  - **Large-scale SDPs.**

- **Lasserre's dense relaxation**
  — theoretical convergence but expensive in practice.
- **The proposed sparse relaxation**
  = Lasserre's relaxation + sparsity
  — no theoretical convergence but very powerful in practice.
- **There remain many issues to be studied further.**
  - **Exploiting sparsity.**
  - **Large-scale SDPs.**

- **sparse SOS and SDP relaxations will work as very powerful methods to compute global optimal solutions of POPs.**

This presentation material is available at

http://www.is.titech.ac.jp/~kojima/talk.html

Thank you!

References

[1] D. Henrion and J. B. Lasserre, "GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi'.

[2] S. Kim, M. Kojima and H. Waki, "Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems". To appear in *SIAM J. on Optimization*.

[3] J. B. Lasserre, "Global optimization with polynomials and the problems of moments", *SIAM J. on Optimization*, 11 (2001) 796–817.

[4] P. A. Parrilo, "Semidefinite programming relaxations for semi algebraic problems". *Math. Prog.*, 96 (2003) 293–320.

[5] S. Prajna, A. Papachristodoulou and P. A. Parrilo, "SOSTOOLS: Sum of Squares Optimization Toolbox for MATLAB – User's Guide".

[6] H. Waki, S. Kim, M. Kojima and M. Muramatsu, "SparsePOP: a Sprase Semidefiite Programming Relaxation of Polynomial Optimization Problems".